FINAL FY03 REPORT FOR

"Naval Automation and Information Management Technology"

(N00014-03-1-0780)

Jerry Pratt, Peter Neuhaus, Jeffrey Bradshaw, Niranjan Suri, James Allen, Lucian Galescu

University of West Florida Institute for Human and Machine Cognition, Pensacola, FL

Submitted to:

Dr. Gary Toth
Office of Naval Research
800 N. Quincy Street
Arlington VA 22217-5660

ihmc
INSTITUTE FOR HUMAN & MACHINE COGNITION

University of West Florida, 40 South Alcaniz Street, Pensacola, FL 32502

I. Summary

Military uses of unmanned systems are growing. The use of unmanned systems, particularly UAVs, in the campaign in Afghanistan and in Iraqi Freedom operations demonstrated beyond any doubt the effectiveness and viability of unmanned systems in ISR as well as weapons delivery missions. As a result, in future military scenarios, large numbers of unmanned ground, air, underwater, and surface vehicles will work together, coordinated by an ever smaller number of human operators. In order to be operationally efficient, effective and useful, these robots must have competent physical and sensing abilities, must be able to perform complex tasks semi-autonomously, must be able to coordinate with each other, and must ultimately be observable and controllable in a useful and intuitive fashion by human operators.

Under the Naval Automation and Information Management Technology Program (NAIMT), The Institute for Human and Machine Cognition (IHMC) of the University of West Florida has conducted advanced research on unmanned systems in the areas of (1) unmanned underwater vehicle mobility, (2) human-agent teamwork and agile computing and (3) mixed initiative human control. Progress made in FY03 in each of these three areas is described below.

II. Unmanned Underwater Vehicle Mobility

Unmanned Underwater Vehicles have many potential applications in Navy mission scenarios. Inspired by dolphins, sea lions, and other swimming animals, we have been investigating the use of biologically inspired propulsion for unmanned underwater vehicles. Our long term goal is to develop next generation UUV platforms with maneuverability, stealth, and efficiency characteristics approaching those of biological systems. These next generation UUVs will employ biologically inspired morphology and propulsion mechanisms and will be more agile and quiet than conventional propeller-driven designs: Turning time will be on the order of several seconds with a near zero turning radius; and the audible signature of the vehicles will be nearly indistinguishable from ecological noise. Due to the advantages of these next generation UUVs, they will find application in many different Navy mission scenarios including mine counter measures, ship and facility protection, and surveillance.

In the first year of this project, we performed concept designs of UUVs inspired by sea turtles, sharks, and dolphins. We decided on a dolphin-inspired design due to the impressive three-dimensional maneuverability of dolphins and also due to geometric design constraints. Using a slightly modified version of already developed, force controllable, Series Elastic Actuators, we have been able to design twelve actuated degrees of freedom into a body approximately the size of a real dolphin.

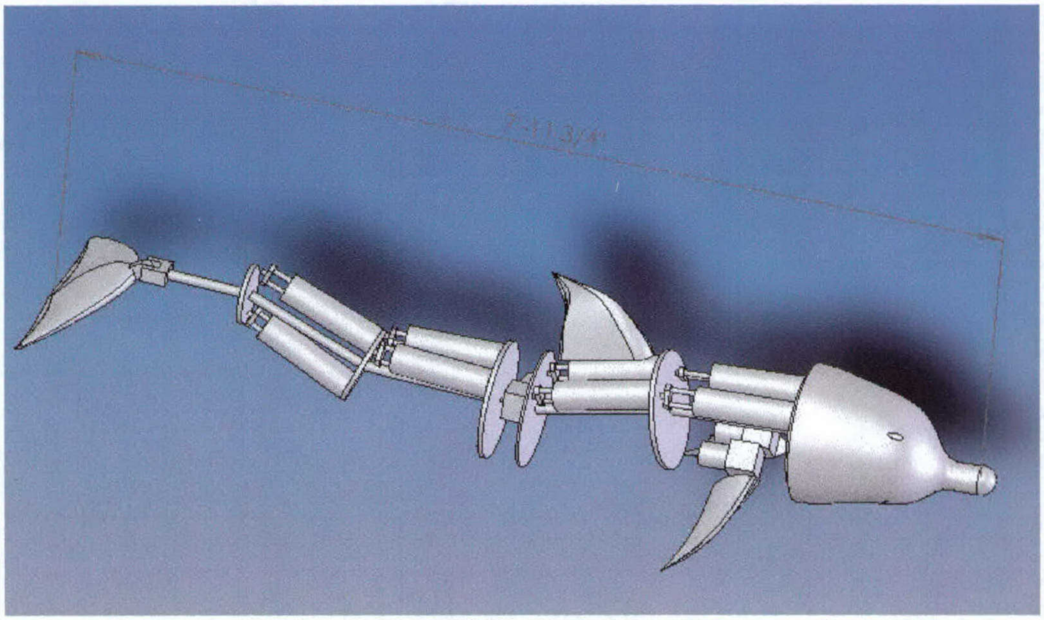
One of our long term goals is to use Computational Fluid Dynamics simulations in the design of both the vehicles' morphology and their control systems. Due to the complex nature of hydrodynamics, most underwater vehicles have been developed in an iterative build-and-test fashion. By using accurate simulations of the vehicles, there is the potential to achieve numerous design iterations before making a single part, thereby improving the capabilities of our designs while using fewer resources. In the first year of this project, we have worked with the hydrodynamics group at NSWC-PC to develop these simulation tools. We have made progress on linking 2D fluid and rigid body simulation engines and simulating a 12 link "Sea Snake" robot. More work is required to make the tools fully 3D and work with complicated body shapes, such as the dolphin-inspired UUV we are building.

Year One Progress

In the first year of this project, we performed a preliminary design of a dolphin-inspired robot, developed coupled computational fluid dynamic and rigid body simulation tools in conjunction with NSWC-PC, investigated potential configurations for underwater swimming exoskeletons, and performed a feasibility analysis of one exoskeleton configuration.

UUV Design

Our UUV design is inspired by a dolphin, particularly the impressive maneuverability of a dolphin due to the flexible spinal cord. Our initial design, shown below, will have 12 degrees of freedom, six in the swimming plane, three for lateral motions, one for twist, and two to rotate the pectoral flippers. The robot will be approximately 8 feet long and employ force controllable Series Elastic Actuators at each of the main swimming joints. The embedded computer will be located in the head, and batteries will be located in the free space in the middle body segments.



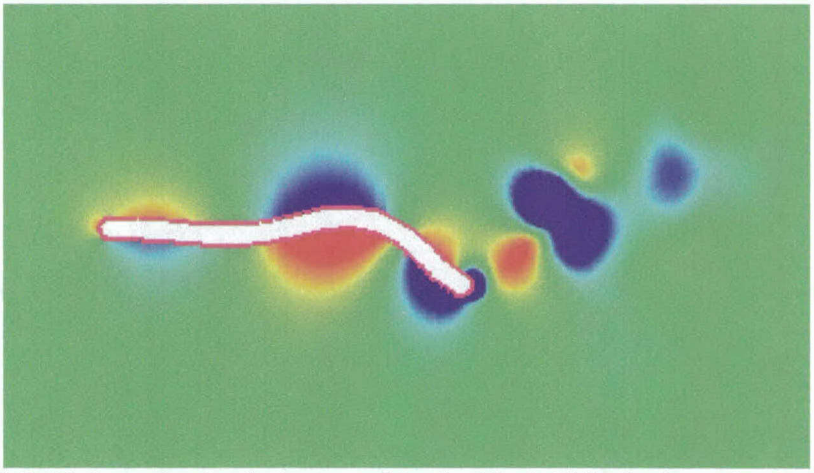
Preliminary design of a dolphin-inspired UUV. The robot will have twelve actuated degrees of freedom and be approximately 8 feet long. Shown are actuator locations. Electronics will be located in the head and batteries will be located in the space in the two middle sections.

UUV Simulation Tools

In conjunction with the Computational Fluid Dynamics Group at Coastal Systems Station, in Year One we have made progress on the development of a simulation tool for accurately simulating shape-changing underwater vehicles. Our long term goal is to develop a tool that will have the following desired features:

- Accurately (in both time and space) simulate high Reynolds number, incompressible, hydrodynamics with moving boundaries.
- Accurately simulate the rigid body dynamics and actuators of the unmanned vehicle.
- Accurately couple the hydrodynamics and rigid body dynamics.
- Easily enter new vehicle morphologies and control systems.
- Interface with mechanical design tools.
- Display hydrodynamic properties and vehicle motion with intuitive displays and controls.
- Run on networked computers so that multiple designs can quickly be analyzed in parallel.

To date we have implemented a 2D version of the simulation tool that simulates coupled rigid-hydrodynamics. As a test case, we have simulated a 12 link "SeaSnake" robot, shown below. The robot has rotational joints with servo models at each joint. A control system that could be used on a real robot has been implemented and used to make the SeaSnake swim at approximately 1 body length per second.



Computational fluid dynamic simulation, combined with rigid body dynamic simulation of a 12 joint, planar, "Sea Snake" robot.

The computational dynamics engine of the simulation has two parts, the hydrodynamic engine and the rigid body engine. The hydrodynamics engine is based on current work being performed by the Computational Fluid Dynamics group at NSWC-PC (Wright and Smith 2001, Smith and Wright 2002, 2003). We are using an edge-based finite volume method for discretizing the unsteady incompressible Navier-Stokes equations using hybrid unstructured meshes. The pressure-velocity coupling procedure we use ensures a divergence-free condition on the velocity field, a condition necessitated by fluid incompressibility. An arbitrary Lagrangian-Eulerian (ALE) form of the fundamental conservation laws allows for arbitrary movement in time of the domain and interior control surface boundaries. These boundaries are adjusted with a simple algebraic grid movement strategy. The hydrodynamic solution is advanced in time using a two-stage implicit Runge Kutta time integration method. The resulting hydrodynamics engine is second-order accurate and geometrically conservative for arbitrary time-dependent meshes.

The rigid body engine is based on the Yobotics! Simulation Construction Set, a rigid body dynamics package developed by the Principal Investigator. This package uses the Featherstone dynamics algorithm (Featherstone 1987) as its computational engine. This algorithm is O(n) in the number of degrees of freedom and easily incorporates both internal joint torques and external forces applied to the object. It has been expanded to incorporate hydrodynamic stresses due to the interaction of the hydrodynamic model.

Coupling between the fluid and rigid body domains occurs at discrete time steps. The vehicle shape and shape derivatives dictate the boundary conditions for the fluid computations. The CFD engine solves for the fluid conditions for the next time step. The stresses and pressures from the fluid are then integrated over small boundary areas to determine interaction forces acting on the robotic platform. Given these interaction forces and the internal actuator forces, the rigid body engine solves for the accelerations of the vehicle. These accelerations then are integrated to determine the vehicle shape and shape derivative for the next time step.

The accuracy and usefulness of this simulation package is being evaluated by using it to simulate existing examples of robotic fish. In particular, we have been starting to compare the results of simulating the MIT RoboTuna with the numerous parameter variations that were performed by Barrett (1996).

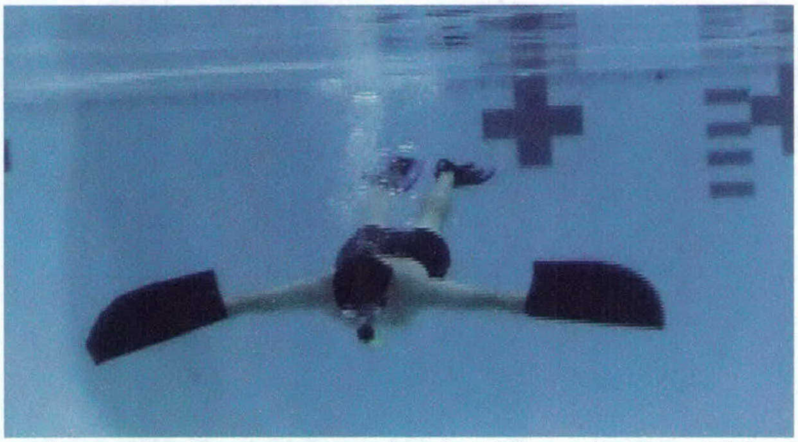
Underwater Exoskeleton Feasibility

The maneuverability, stealth, and efficiency gains of biologically inspired propulsion may also be useful for Navy divers. As part of the Year One work, we investigated the feasibility of an underwater propulsion exoskeleton that can be used by divers to increase their swimming speed and maneuverability, while maintaining their stealth.

Our exoskeleton feasibility analysis has shown that we can build an underwater exoskeleton that will help propel a diver at two knots for up to six hours with 15 kg of battery. Encouraged by this analysis, we have developed a concept for PISCES: Performance Improving Self-Contained Exoskeleton for underwater Swimming, shown below. PISCES would have actuators at the user's hip and knees with battery packs stored on the users back. It could potentially double the range of Navy SEALs, while retaining their stealth.

Concept design for a Performance Improving Self-Contained Exoskeleton for underwater Swimming (PISCES). Four actuators are shown driving the user's hips and knees. Canisters on the user's back contain enough battery power for an estimated 6 hours of swimming at 2 knots.

Inspired by the way that sea turtles and penguins swim by essentially flying through the water, we also investigated the potential for wing based propulsion. Below is shown one of the "Aquawing" prototypes we made and tested. The Aquawings were promising in that they allowed for swimming faster than one can swim underwater with hands alone. However, they were still slower than swimming with fins. Motorizing the wings with an exoskeleton could potentially result in an entertainment device.



"Aquawing" mock-ups developed to test the feasibility of using lift-based wing-driven exoskeleton to enhance human swimming.

Resulting Papers

Peter D. Neuhaus, Michael O'Sullivan, David Eaton, John Carff, and Jerry E. Pratt, (2004). "Concept Designs for Underwater Swimming Exoskeletons". Proceedings of the 2004 IEEE International Conference on Robotics and Automation, New Orleans, LA, 4893-4898.

References

Barrett, D., Grosenbaugh, M., and Triantafyllou, M. (1996). "The optimal control of a flexible hull robotic undersea vehicle propelled by an oscillating foil." Proceedings of the 1996 Symposium on Autonomous Underwater Vehicle Technology, 1-9.

Featherstone, R. (1987). Robot dynamics algorithms, Kluwer, Boston.

Smith, R. and Wright, J. (2003). "An Implicit Second-Order ALE Method for the Incompressible Navier-Stokes Equations." (Under Review for Numerical Methods in Fluids).

III. Human-Agent Teamwork and Agile Computing

In the years ahead, unmanned systems will be used on an ever-increasing scale [15]. A key requirement for such systems is for real-time cooperation with people and with other autonomous systems. While these heterogeneous cooperating platforms may operate at different levels of sophistication and with dynamically varying degrees of autonomy, they will require some common means of representing and appropriately participating in joint tasks. Just as important, developers of such systems will need tools and methodologies to assure that such systems will work together reliably and safely, even when they are designed independently.

An equally challenging problem involves the fact that unmanned vehicles are subject to communication constraints that limit bandwidth and increase latency. In addition, network disconnection is a concern, whether due to vehicles moving out of communications range, communications being obstructed by terrain, or a tactical need to minimize signal transmissions. Finally, communication may sometimes depend on peer-to-peer networks, where one vehicle communicates with another vehicle by using a third vehicle as a relay. These problems are particularly acute for undersea and surf-zone environments.

Both the dynamics of human-autonomous system coordination and the ongoing management of real-time operational constraints can be addressed by the use of software agent technology. Software agents are loosely-coupled components designed with a variety of built-in communicative and collaborative capabilities. In addition to these built-in generic capabilities, each agent usually serves as a package for some more specific intelligent functionality (e.g., sensing, fusion, analytic, or navigation behavior). The combination of these generic and agent-specific capabilities help enable unmanned vehicles to function as effective "team members" with each other and with other autonomous systems. Under the strict control of administrator-defined policies, one or more software agents may be permitted to populate a given hardware vehicle platform or to move around the network as needed under their own power, operating in dynamically optimized onboard or off-board combinations.

The combination of human-agent teamwork and agile computing capabilities afford a degree of flexibility and responsiveness in the configuration and tasking of unmanned vehicles that goes far beyond what is possible with today's technology. Different tasks and missions place different requirements on the unmanned vehicles and, given their limited processing and storage capabilities, the necessary algorithms for responding to dynamically changing conditions will often need to be pushed to the vehicles in real time. Changing conditions may require adaptive task allocation among humans and machines, including a requirement that other nearby resources may need to be rapidly discovered for immediate exploitation. If a human team member becomes disabled or a vehicle is suddenly destroyed, the survivability of the system depends directly on being able to quickly shift tasks and capabilities among people and platforms consistent with pre-approved operating policies and procedures.

In this research focus, we are addressing these issues through the development of policy-based human-agent teamwork and agile computing infrastructures. These developments will result in a robust teamwork-aware computational infrastructure for unmanned systems that is secure, reliable, and capable.

Year 1 accomplishments.

- **Task T2-1-1: Obligation policy support in KAoS.** We have implemented an initial version of support in KAoS for obligation policies (i.e., constraints that require or waive requirements for certain kinds of actions in a given context), with enablers to enforce them and enhancements to the KPAT user interface to define them. We have also developed an initial implementation of KAoS Robot interfaces, and have begun development of simulation capabilities and viewers. We have converted KAoS policy representations from DAML to OWL.
- **Task T2-1-2: Safe and secure autonomous operation.** We incorporated an initial KAoS policy enforcement mechanism into the agile computing bandwidth management component and have demonstrated an initial version of these capabilities. We have purchased robotic hardware (in consultation with USF so that we have compatible setups) and have established an initial UAV/UGV robotic testbed at IHMC. We have begun integration of our components with USF's in collaobration on the distributed field robot architecture.
- **Task T2-1-3: Effective and natural human-agent interaction.** We performed an initial study of what display and behavior options people find most effective for robots to communicate common states and actions. We developed an initial set of ontologies and notification and event policy implementation components. We have begun to integrate KAoS with the dialogue system components and have demonstrated these capabilities in conjunction with physical robots.
- **Task T2-2: Agile computing.** We have implemented initial version of FlexFeed middleware to provide efficient sensor data feeds. We demonstrated resource discovery, bandwidth optimization, and policy enforcement at the August Review Meeting in Pensacola. Now, we have tarted working on incorporating a Just-in-Time Compiler for Aroma VM, as well as a prototype for location tracking for robotic platforms. We have also demonstrated the notion of proactively directing system behavior based on resource requirements.

Follow on Projects

In follow on projects to this research focus, we will continue development a joint distributed software architecture for cooperative vehicles to fully integrate the human-agent teamwork, agile computing, distributed sensor fusion and control and mixed-initiative human control components. This architecture will be integrated into the overall distributed field robot architecture and will be based on results of investigations in the proposed research, including collaboration with the NSWC-PC cooperative behavior research group. In particular, we will be interested in testing the value of these technologies in reducing demands on the operator and in providing more robust, survivable, and flexible platform behavior.

Resulting Presentations and Publications

Boy, G. & Bradshaw, J. M. (2003). From direct manipulation to agent management. Tutorial presented at the European Cognitive Science Conference (EuroCogSci). Osnabrueck, Germany, 9 September.

Boy, G. & Bradshaw, J. M. (2004). Interaction design of highly automated domain-specific systems. Tutorial presented at the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2004), Vienna, Austria, April 26.

Bradshaw, J. M. & Suri, N. (2004) Policy-based approaches to increasing the acceptability of intelligent systems. Invited seminar at the IBM T. J. Watson Research Laboratory, Hawthorne, NY, 18 February 2004.

Bradshaw, J. M. (2003) A policy-based approach for collaboration in human-agent teams: Issues and applications. Invited keynote speech for the IAT 2003 workshop on Collaboration Agents: Autonomous Agents for Collaborative Environments (COLA '03). Halifax, Nova Scotia, 13 October.

Bradshaw, J. M. (2004) Dimensions of adjustable autonomy and mixed-initiative interaction. IHMC seminar, 20 January.

Bradshaw, J. M. (2004). Bridging the gap between multi-agent, multi-robot, and multi-human teams. *Presentation at the AAAI Spring Symposium,* Stanford, CA: AAAI Press, March 22-24.

Bradshaw, J. M. (2004). Dimensions of adjustable autonomy and mixed-initiative interaction. Department of Computer Science, Brigham Young University, 13 February 2004.

Bradshaw, J. M. (2004). Dimensions of adjustable autonomy and mixed-initiative interaction. Invited colloquium at SRI, International, Palo Alto, CA, June 8.

Bradshaw, J. M. (2004). Semantically-rich policy representation, reasoning, and enforcement. Invited keynote address. Second International Conference on Trust Management. Oxford, England, 29 March-1 April.

Bradshaw, J. M., Uszok, A. & Jeffers, R. (2004). Toward semantically-rich policy representation: Issues and applications. *Conference on the Human Impact and Application of Autonomic Computing Systems (CHIACS2).* Yorktown Heights, NY: IBM T.J. Watson Research Center, 21 April.

Bradshaw, J.M. (2003) A glimpse at the future of agent technology. Invited tutorial at the 2003 IEEE/WIC International Conference on Intelligent Agent Technology (IAT 2003), Halifax, Nova Scotia, 13 October.

Bradshaw, J.M. (2003) Living with agents: From human-agent teamwork to cognitive prostheses. Invited colloquium at University of Southern California Information Sciences Institute, Marina del Rey, CA, 3 December.

Bradshaw, J.M., Jung, H., Kulkarni, S., & Taysom, W. (2004). Dimensions of adjustable autonomy and mixed-initiative interaction. In M. Nickles, M. Rovatsos & G. Weiss (Eds.), *Agents and Computational Autonomy: Potential, Risks, and Solutions*. Lecture Notes in Computer Science, Vol. 2969. Berlin: Springer-Verlag, pp. 17-39.

Bradshaw, Jeffrey M. (2003). Living with agents: From human-agent teamwork to cognitive prostheses. Inaugural lecture as honorary visiting researcher at the Centre for Intelligent Systems and their Applications and AIAI at the University of Edinburgh, Scotland, 4 September.

Bradshaw, Jeffrey M. (2003). Living with agents: From human-agent teamwork to cognitive prostheses. Invited lecture, Siemens AG Corporate Technology, Munich, Germany, 12 September.

Feltovich, P. J., Bradshaw, J. M., Jeffers, R. & Uszok, A. (2003) Social order and adaptability in animal, human, and agent communities. *Proceedings of the Fourth International Workshop on Engineering Societies in the Agents World*, Imperial College, London, 29-31 October.

Feltovich, P. J., Bradshaw, J. M., Jeffers, R., Suri, N., & Uszok, A. (2004). Social order and adaptability in animal and human cultures as analogues for agent communities: Toward a policy-based approach. In A. Omicini, P. Petta, & J. Pitt (Eds.) *Engineering Societies in the Agents World IV*. Lecture Notes in Computer Science Series. Berlin, Germany: Springer-Verlag, pp. 21-48.

Jung, H., Bradshaw, J. M., Kulkarni, S., Breedy, M., Bunch, L, Feltovich, P., Jeffers, R., Johnson, M., Lott, J., Suri, N., Taysom, W., Tonti, G., & Uszok, A. (2004). An ontology-based representation for policy-governed adjustable autonomy. *Proceedings of the AAAI Spring Symposium,* Stanford, CA: AAAI Press, March 22-24.

Tonti, G., Bradshaw, J. M., Jeffers, R., Montanari, R., Suri, N., & Uszok, A. (2003). Semantic Web languages for policy representation and reasoning: A comparison of KAoS, Rei, and Ponder. *Proceedings of the International Semantic Web Conference (ISWC 03).* Sanibel Island, Florida, October.

Uszok, A., Bradshaw, J. M., Hayes, P., Jeffers, R., Johnson, M., Kulkarni, S., Breedy, M. R., Lott, J., & Bunch, L. (2003). DAML reality check: A case study of KAoS domain and policy services. Internal technical report. Short version of this paper appears in *Proceedings of the International Semantic Web Conference (ISWC 03).* Sanibel Island, Florida, October.

Uszok, A., Bradshaw, J. M., Jeffers, R., Johnson, M., Tate, A., Dalton, J., & Aitken, S. (2004). Policy and contract management for semantic web services. *Proceedings of the AAAI Spring Symposium,* Stanford, CA: AAAI Press, March 22-24.

IV. Mixed-Initiative Human Control

In addition to the problem of having a group of unmanned vehicles coordinate their activity with each other, we also need to integrate humans into the coordinated activity in an intuitive and natural way. It is essential that the human controllers can understand and assess the ongoing situation as it evolves, and modify the autonomous systems' activities as needed. In the long term, the addition of spoken language to the usual graphical modalities is considered to be one of the most promising ways to provide the natural, intuitive, and efficient interface needed to achieve these goals. In addition, in some applications spoken language is the only option: when the hands or the eyes are not available (e.g., because of simultaneously operating a vehicle), when interacting with devices that have limited graphical abilities (e.g., hand-held devices), or when interacting directly with wall displays (e.g., the kind designed for the Naval Warfare Assessment Division). Although full understanding of spontaneous spoken language is far from being a solved problem, we are making substantial progress towards developing useful systems that use speech in some limited form.

In this project focused mainly on synergistically combining spoken language with rich graphical interfaces. Our approach, however, is general in scope, and can be easily applied to configurations with less weight on the GUI. We developed the system within the collaborative framework envisioned under Research Focus R2. This involved keeping track of the dialogue context and the problem-solving context in order to facilitate economy of expression (e.g., using pronominal references), clarifications, questions, and multi-unit input (e.g., complex commands) across different input modalities.

As desirable as it is, a fully unconstrained dialogue system that supports true cooperative behavior is currently beyond the state of the art. We focused this research on problems with long-term benefits that also enable more limited practical systems in the short term. From a technical point of view, a key concern was developing an architecture that provides humans with intuitive and flexible control of the unmanned systems. This architecture must support:

- the ability to use contextual and linguistic constraints to enhance the recognition of spontaneous speech;
- the intuitive presentation of information with a capability to "drill-down" and present information in different ways in response to questions;
- the collaborative development of plans in which the humans and the unmanned vehicles combine their knowledge and capabilities to develop the most effective course of action in response to situations;
- the tasking and collaborative re-tasking that must occur as the situation evolves; and
- the explicit discussion and negotiation of responsibilities to define the parameters for adjustable autonomy.

Year One Progress

During the first year of the project we completed the integration of a version of the TRIPS dialogue system with the KAoS framework, which enabled us to demonstrate robust and effective participation of a human user working with a team of real robots on a simple mine-finding task.

TRIPS-KAoS Integration – We realized early on that moving swiftly towards integrating our multi-modal dialogue system with the KAoS framework (Research Focus R2) would benefit both groups. To this end, we designed and built an initial TRIPS/KAoS interface module that connects the two systems. TRIPS is able to convey the user's requests, queries, etc., to KAoS and, conversely, to receive, via KAoS and FlexFeed, information from the robots (e.g., their location, video streams, etc.).

In addition, we implemented an initial mechanism for mapping between the KAoS ontology and the TRIPS ontology. We developed an ontology mapping language and built a system that uses the mappings to transform from the language-based logical form in TRIPS and the KAoS Ontology. This is a general mechanism. To handle a different domain/ontology, we simply have to define a new set of mapping rules.

Of particular interest to us is the possibility for the user to discuss policies that affect the performability of various actions, or the manner they are performed. We implemented an initial mechanism for manipulating authorization policies on robots (e.g., giving them authorization to move). Again, we intend to continue working with the Human-Agent Teamwork group towards new interaction models to drive the dialogue regarding authorizations and permissions.

Speech Recognition and Language Modeling – In language processing, the challenge and our primary focus is spoken language understanding, not speech recognition. However, poor speech recognition is a serious impediment towards achieving good performance on the understanding task. Resorting to fixed sets of pre-defined commands and queries is not an option because it would both severely affect the naturalness of the interaction, and be inadequate in a complex domain. We, therefore, decided to use statistical language models to guide the speech recognition process, which allow for a wide coverage of potential spoken language input. However, this flexibility comes at a cost: statistical models have to be trained on fairly large amounts of data, which in the case of new domains simply don't exist. Moreover, collecting training data is a very costly enterprise.

For this project we used and refined a technique that we developed initially for the TRIPS-Pacifica domain (Galescu, Ringger & Allen, 1998). The idea is that, for practical dialogue, linguistic input is relatively tightly associated with the capabilities of the system and the features of the application domain. For example, in our domain we have robots that are movable, can detect objects, transmit pictures or video, etc. This information is relatively quickly put in the form of a small set of utterances that are then generalized into a grammar; this grammar can be used to generate bigrams (word pairs) for training a statistical language model. The approach dramatically reduces the time and effort to deploy a speech recognition component for a dialogue system. Importantly, it leads to good speech recognition performance, and excellent language model portability[1] and adaptability[2].

Language Understanding – In this area, much of the work during the first year of this project was directed towards improving the architecture of the TRIPS system to allow the sharing of the ontology and the lexicon across the various modules that use them: speech recognition, parsing, language generation. This will greatly benefit future system development by providing a single point of entry for maintenance and learning of new concepts and words.

Multi-Modal Generation – The goal of the generation component is to convey meaning and intent to the user in a clear and concise manner. If the system is to convey information through language, naturalness is also an important factor; for example, the phrase "an interesting object" sounds more natural than the semantically equivalent "an object that is interesting". Providing natural generation through both language and graphical modalities poses an additional challenge: for example, sometimes it might be enough to flash a symbol on a map, while other times it would be important to draw the user's attention by speaking "The object is here!" in addition to flashing the symbol.

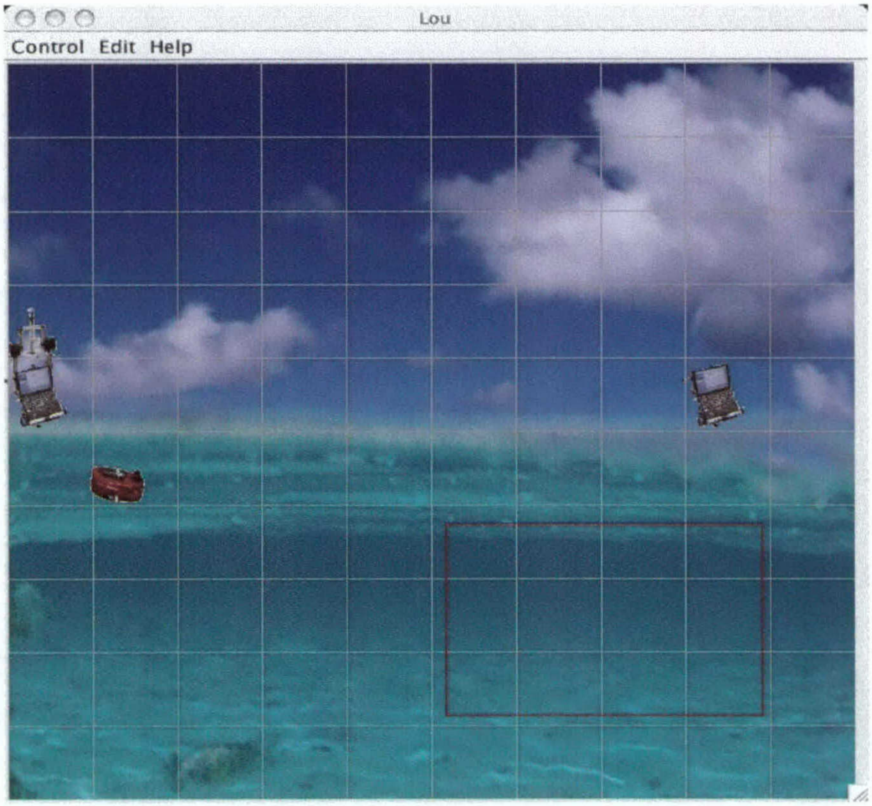
In year one we developed a two-stage generic natural language generation technique that first over-generates a set of candidate utterances from a semantic representation, and then uses a statistical language model similar to the one used for speech recognition to decide which of the candidate utterances is most appropriate (Chambers & Allen, 2004). The first stage is based on a general purpose grammar of English sentences, while the second phase uses a mixture of domain-independent and domain-specific statistical language models; the heavy use of domain-independent linguistic information leads to increased portability of this component and robustness in unexpected situations. Whereas this first stage may generate utterances that are incorrect or inappropriate, the second stage tries to eliminate them by making lexical and phrasal choices as well as an overall utterance style that are most appropriate for the domain at hand. Moreover, we started to explore the possibility that this second stage be made more adaptive in order to match the user's lexical choices and speaking style for increased naturalness and communication effectiveness.

In additional to language output, the generation component also has capabilities to generate communicative acts using graphic modalities and combining them with speech (as in the example given above involving flashing objects on the map to indicate the location of a robot or other object in the field).

---

[1] For example, adding words or commands in connection when a new robot, with new capabilities becomes available would require only slight additions to the grammar. In general, even when changing the domain completely, parts of the grammar may be re-used.

[2] That is, as the dialogue system starts being used, data can be collected and used to adapt the initial language model for increased recognition quality.



Active map, allowing the user to assess the situation; both the user and the system can combine speech with GUI events on the map

Multi-Modal Displays – For the current domain we have implemented an active map that provides the user with a representation of the terrain and the location of the robots and the mines being discovered. In addition to being used for output by the generation component, as discussed above, the map is also used for input. For example, the user may select an area or object of interest, and use pronominal and deictic references to them (e.g., "Search *this* area", or "What is *this robot* doing?" or even "What is *this one* doing?"). These references are linked contextually to the GUI events to arrive at the correct interpretation of the user's utterance.

Graphical information from the robots (e.g., pictures or video) can also be presented to the user, who can manipulate certain features of the respective displays through language. For example, the user may request an increase in the resolution of a video stream. Currently, however, it is not possible for the user to refer to the contents of such a display; for example, if the human user spots some letters on the body of the mine, he cannot ask the robot "Can you zoom in on these letters here," but he could direct the robot to place itself and its camera in an advantageous position and zoom in.

Simulation environment – We started to develop a simulation environment that will model UUVs and the environment in a physically correct manner, enabling us to carry out experiments with the full system on an ongoing basis.

Resulting Papers

Chambers, N., & Allen, J. (2004). Stochastic Language Generation in a Dialogue System: Toward a Domain Independent Generator. To appear in the 5th SIGdial Workshop on Discourse and Dialogue. Boston, USA.

References

Allen, J., Byron, D., Dzikovska, M., Ferguson, G., Galescu, L., & Stent, A. (2000). An Architecture for a Generic Dialogue Shell. Journal of Natural Language Engineering, 6(3), 1-16.

Allen, J. F., & Perrault, C. R. (1980). Analyzing Intention in Utterances. Artificial Intelligence, 15(3).

Cohen, P. R., & H. J. Levesque (1980). Speech Acts and the Recognition of Shared Plans. Paper presented at the 3rd Biennial Conference of the Canadian Society for Computational Studies of Intelligence, Victoria, B. C.

Ferguson, G., & Allen, J. (1998). TRIPS: An Integrated Intelligent Problem-Solving Assistant. Paper presented at the NCAI (AAAI-98), Madison, WI.

Galescu, L., Ringger, E. & Allen, J. (1998) Rapid Language Model Development for New Task Domains. Proceedings of the First International Conference on Language Resources and Evaluation (LREC), Granada, Spain.

V. SUMMARY

Substantial progress has been made in all three areas of research undertaken in the first period of performance for this effort in Naval Automation and Information Management Technology. A follow-on year has been funded to continue this promising work.